Solid Earth Discuss.,
https://doi.org/10.5194/se-2018-77-AC2, 2019

Solid Earth
Discussions
Open Access

SED

Interactive comment

Interactive comment on "Precision of continuous GPS velocities from statistical analysis of synthetic time series" *by* Christine Masson et al.

Christine Masson et al.

masson@gm.univ-montp2.fr

Received and published: 23 January 2019

General Comments: Masson et al. present an analysis of constructed position time series that were made using analytical forms of signal and noise that are typically observed in geodetic GPS data. They then use these synthetic time series to identify the factors that contribute the most to uncertainty in the estimated trend. Once the most important factors are identified (time series length, spectral content and amplitude of colored noise, etc.), they offer a few rules of thumb that can be applied to categorize time series according to how precise they are expected to be. One of their conclusions is that the time series duration is invariably the most influential factor in maintaining low uncertainty in velocity estimation, which is very important when considering how GPS networks are funded and maintained. Aside from the focus on synthetic time se-

Printer-friendly version

Discussion paper

ries, two elements in the paper stand out as looking new to me. First, they employ a regression-tree approach that rank orders parameters in terms of their overall impact on the velocity uncertainty. This is useful since these factors are sometimes known for real data in advance and can be used to generate expectations for which time series provide the lowest uncertainties, before any more detailed analysis is undertaken. Second they introduce a new method for scanning the time series for discontinuities that are undocumented, i.e., their existence and time of occurrence are unknown beforehand. Their method is interesting because they flip the problem by determining which epochs contain *no* step to within data uncertainty, thereby narrowing the set of epochs that could have steps. I had a number of suggestions, mostly minor which I placed in the technical comments below. The introduction could benefit from some short additional text, possibly in the last paragraph, on the general value of looking at synthetic time series as opposed to 1) real ones when so many are available, or 2) simple formulas that mathematically represent the content of signals+noise in them (e.g., Williams, 2003). The answer might be e.g., the ability to know the true answer in order to evaluate the validity of false and true detections, which might be obvious at the onset to expert readers but not everyone. That paragraph would be a good place to also mention the limitations of an analysis like this, since many real GPS time series contain signals of types not included in their synthetic tests. They mention a few examples in the paper but do not discuss the impact of the potential presence of these signals in detail.

> We added the following elements to address this point:

Page 1 Line 27: ... However, several state-of-the-art applications of ... be defined with increasingly better precisions, ... Page 2 Line 23: In this study, we estimate the potential precision of GPS velocities through a statistical analysis of synthetic position time series that are representative of standard GPS data. We focus on continuous time series with a daily sampling frequency (i.e., permanent rather than campaign mode) to test the effect of colored noise, periodic signals, and position offsets (with a new method for automatic offset detection). The use of synthetic data allows a detailed analysis of the velocity estimations compared to the target ("true") velocities and of the specific contribution of each parameter that can be treated independently. On the contrary, such an analysis would not be possible with real GPS data in which the true value and role of each parameter cannot be fully deconvolved. The parameter range used in the synthetic data is representative of typical average data and excludes the potential effect of transient phenomenon, such as slow slip or postseismic events, or that of pluri-annual hydrological processes. The impact of such phenomenon is addressed in several recent studies (e.g. Altamimi et al., 2016; Chanard et al., 2018) and could be included in more detailed synthetic analyses beyond our present study. We illustrate ...

Some detailed/technical comments: Page 2 line 26. Could replace "low deformation" with "low rate of deformation"

> We made the necessary changes. Thank you. (Page 2 Line 32)

Page 2 line 31. In equation 1 they may have meant to use H(t) rather than the Kronecker delta function to indicate the occurrence of a step in the time series. H(t), the Heaviside function, is zero before t and one after t, and is also the time integral of the Dirac delta function. The Kronecker delta function is a discrete version of the Dirac delta function, in physics literature. Here Masson et al., define delta(t) in a way that works for their paper so it is probably all OK and self-consistent here, but might cause some minor confusion to call it the "Kronecker delta function".

> Yes indeed, changed that. Thank you for pointing this mistake.

Figure 2. It seems odd to me at first to lump all the horizontal and vertical data together in the analysis, and in this one plot. I guess it all works out in the end. But I wondered if including a new binary parameter in the regression tree, horizontal vs. vertical time series, would have a strong predictive ability in the tree.

> The main difference between horizontal and vertical time series is the noise level. The ranges of values overlap between horizontal and vertical noise levels, so we prefer not to make a distinction and to carry out the analysis without testing for "horizontal series vs. vertical series". As the reviewer indicates, this “works out in the end” since the noise levels acts as an indicator that separates the first-order horizontal and vertical data, without forcing an a priori distinction, which is quite remarkable.

Figure 3. The caption lists values of k as positive, but on page 3 they are said to be always negative.

> Indeed, we made the necessary changes. Thank you.

Page 5 line 26. It is a little confusing sometimes that they interchange the terms “accuracy” and “uncertainty”. For example, Figure 4 is a nice plot, but “accuracy” should be changed to “uncertainty” since accuracy should improve (increase) with time series duration, but the quantity shown decreases with time series duration.

> We agree that this is confusing... We used the term “accuracy” because it corresponds to what we measure, i.e. the deviation of the estimated velocity from the true velocity. Hence a “high accuracy” in common phrasing actually corresponds to a small number (small deviation from the true value). Using “uncertainty” would help (i.e., “high uncertainty” = large number), but it leads to the confusing situation pointed out by the reviewer in which both terms are used but are not interchangeable. In the original manuscript, we also used the term “precision” as a generic word to be more in line with classical studies of GPS velocity uncertainties, which discuss “precision” (dispersion around a central value) and not “accuracy” (which cannot be known for actual GPS data). In order to clarify this, we propose to replace the term “accuracy” by “bias”, which corresponds to the same concept (deviation from true value) but has an inverse amplitude, i.e. “high bias” = large number. Thus in Figure 4 (and others), the increase in the time series duration leads to a decrease in the bias associated with a smaller number. We also added a couple of sentences to explain this (cf. P2L23, comment

above).

We added the following elements to address this point:

Page 3 Line 1: Velocity bias – For each time series, the calculated velocity is compared with the true (imposed) velocity. The absolute value of the difference between the two is termed “velocity bias” and represents the deviation of the calculated velocity compared to the truth. We choose the term “bias” rather than “accuracy” in order to avoid confusion (e.g., a high accuracy associated with a small number) and different definitions of “accuracy”. For each analysis, the velocity bias distribution is characterized by two statistical estimators: 95% confidence limit (noted v95) – This estimator is the 95% quantile of the bias distribution and represents a 95% confidence in the estimated velocities. Probability of 0.1 mm yr-1 (noted p01) – This estimator is the percentile associate with a velocity bias of 0.1 mm yr-1. E.g., p01 = 75% indicates that a 75% probability that the velocity bias be smaller than or equal to 0.1 mm yr-1. Precision – We limit the usage of the term “precision” to the general concept of “quality” of a velocity estimation, regardless of its origin and whether it corresponds to a systematic error (bias) or a measurement repeatability (dispersion). Standard error and Uncertainty – For each time series, the calculated velocity and other parameters are associated with standard errors estimated as part of the linear inversion (cf. Section 3). These standard errors are used as estimators of the uncertainty on each calculated velocity.

Also in Figure 4, they should state in the caption what is the meaning of the vertical extent of the vertical black bar, and also what is indicated by the extent of the blue box. Then in 5, 6, 7 it can be said they are as in Figure 4. These plots may be standard in some literature but probably not everyone will already know the details of construction.

Caption of Figure 4: . . . noise. Whiskers diagrams show the data quartiles (25, 50, 75%) in blue, the extremes (0%, 100%) with the vertical black line, and the 95 percentile (v95) with the horizontal black line.

Page 5 line 25. “possible asymptotic value ca. v95 = 0.05 mm yr-1”. Possibly? In

Figure 4 it looks like the uncertainty is still decreasing, though more slowly, at duration 20 years. I would have thought that theoretically the asymptote would be v95=0 and maybe we will do at least a little better if we run a GPS station for 100 or 1000 years. I don't see evidence of an asymptote at 0.05 mm yr-1.

> Yes indeed, the use of the word asymptote is unjustified. You are absolutely right.

Page 6 Line 12: For series longer than 15 years, all v95 are smaller than 0.1 mm yr-1. A near-exponential decrease of v95 is observed as a function of the duration of the series with a sharp slowdown from 15 years of data.

Figure 8. I think in the caption v should be v_95? The "95" is dropped in several places when it should be included.

> No, it is actually the velocity bias of each individual time series (and not the 95 percentile of the full dataset). This has been clarified in the caption:

. . . Distribution of the ratio of the velocity bias to its standard error for each individual time series.

Page 6 line 18. Probably meant "if the series is short."?

> Yes, exactly. Thank you.

Page 7 Line 8: This effect is more important if the series is short.

Page 8 line 11. it would be better to use a lower case t, rather than T for the step time so as not to confuse with time series duration.

> Yes, thank you for this idea. In the text we have replaced Ti with ti.

Page 8 line 14. "consists in" should be "consist of"

> Yes, exactly. Thank you. (Page 9 Line 6)

Page 8 line 16. Instead of "amp_off", notation for non-offsets might be better stated in same class as true offsets, e.g., C_notanoffset or something shorter.

> The notation was confusing. We clarify it to use the same notation as the time series equation (eq. 1) to show that we are testing the artificial offsets added:

Page 9 Line 8: . . . The series is then inverted to estimate all offset amplitudes (Ci) and their associated standard errors (σci) jointly with the other model parameters (velocity, seasonal signal, etc.). The offset with the smallest amplitude (CS) is then identified and a simple significance test is performed:

$|C\_S| \geq b .\sigma\_cs$ (5)

If the amplitude (CS) is larger than its scaled standard error (b.σcs), the offset is considered significant. Because the test is performed on the smallest offset and the offset standard errors are similar in the majority of cases, we then consider that all offsets are significant and we keep them in the model. In the opposite case, the smallest offset is rejected and the inversion is redone with the remaining offsets in order to test the new smallest offset, until a significant offset is found or none remains.

Page 8 line 25. I have a few questions about their interesting new offset detection method. First, it seemed that a part of the explanation may be missing. It is stated that it is repeated “until a significant offset is found”. But once an offset is found there could be others

> True. The objective is to remove all non-significant offsets. The assumption is that if the offset with the smallest amplitude is found to be significant, then all others are significant (because they all have similar standard errors), and thus the tests can be stopped.

and, since only evenly spaced arrays of offsets are tested in each iteration, there may be epochs that have not yet been tested. So how does the algorithm guarantee the completeness of the scan for a step at every epoch?

> All possible epochs are not tested. We only test for potential offsets at fixed epochs (every 20 days). The assumption is that a real offset at any given epoch will be caught

by the forced artificial offset located less than 10 days directly before or after. As such, we do not find the exact date of the real offset but its approximate date +/- 10 days. This method cannot resolve real offsets situated within a few (10-20) days of each other. They will be lumped into a single artificial offset, but we assume that its effect on the estimated velocity will be a good proxy of the combined effect of the real offsets.

Secondly, if a large true step exists and the adjacent epoch is tested, it will likely be evaluated as a significant step. Is there a mechanism to replace the adjacent epoch with the correct one once it has been tested?

> No, this is not included.

When the process is repeated are steps and non-steps identified in previous iterations excluded from being considered as steps? If so I expect that would improve efficiency and reduce ambiguity in the algorithm.

> At each iteration, the fixed date of the non-significant offset found in the previous iteration is removed. All other remaining dates are kept, offsets at these dates are resolved, and the smallest is tested for significance.

Finally, could this method be applied to real data? It seems that the calibration method determining b and delta_t in Appendix B relies on the quantity of false and true identifications, so might not be available for real data(?).

> Indeed, the calibration is not possible on real data. Altogether, it is important to keep in mind that this method was only developed as a simple and quick way to test the impact of offsets and their resolution of the velocity estimations. Our tests on synthetic data allow us to show that the method works, statistically as well as others, but we did not try to fine-tune or improve it.

Page 9 line 21. “real” need not be in quotes.

> Indeed, we made the necessary changes. Thank you.

Page 9 line 27. “is no the case” should be “is not the case”

> Indeed, we made the necessary changes. Thank you again. (P10L17)

Page 10 line 14-22. These classifications may be useful but possibly a bit dismissive of the utility of some of the categories when signals are large enough to stand out from the noise. For example the Oregon coast rises >4 mm/yr owing to elastic strain accumulation on the subduction zone. Even short time series may be useful there.

> True. Our implicit purpose here was to classify the GPS velocities for applications that require sub mm yr-1 precision. We rephrased this section better explain this and not imply a generic application of this classification.

Page 1 Line 15: . . . less than 4.5 years are not suitable for studies that require sub mm yr-1 precisions; (3) Series of intermediate . . . Page 11 Line 4: . . . that may be applicable to actual GPS data used for high-precision (sub mm yr-1) studies, considering the fact that series duration is the key parameter . . . Page 14 Line 20: . . . cannot be used for application that require a precision better than 1.0 mm yr-1, except . . .

Page 10 line 31. “These results may indicate a lower limit in velocity accuracy ca. 0.1 mm yr-1”. But it said in the previous sentence that some were 0.05 mm yr-1...

> True. The ambiguity stems from the fact that 0.05 mm yr-1 is associated with noise-alone series (no offset), whereas the following sentence gives a more general value of 0.1 mm yr-1 derived from the generic cases (noise + offsets). This is clarified by removing the number:

Page 11 Line 20: . . . noise effect as a function of time stagnates ca. 15 to 21 years (cf. Fig. 4 and section 3.1). Our results may indicate an overall lower limit on the velocity bias ca. 0.1 mm yr-1 due . . .

Page 11 line 13. “v” should be “v_95”? Also “A ratio of 1 corresponds to a standard error equal to its velocity”. In a Gaussian distribution +/- one standard deviation contains 68% of the samples, whereas the definition of v_95 in this paper is the limit that contains 95%. So would not a ratio of 2 indicate that the standard error and velocity accuracy are similar?

> This section (and the associated figure caption) was not clear. We do not compute the ratio of v95 over a standard error (because v95 is the 95% quantile of the whole dataset), but rather the ratio of each estimated velocity over its standard error (for each individual time series). Thus, ratios of 1 and 2 should correspond to 68% and 95% of the populations. We clarify this in the text:

Page 12 Line 3: . . . We can test the robustness of these standard errors in comparison with their associated velocity biases by computing the ratio of the velocity bias to its standard error for each individual time series. A ratio of 1 corresponds to a standard error equal to its velocity bias; a ratio smaller (greater) than 1 corresponds to a standard error greater (smaller) than its velocity bias. Owing to our stochastic approach, and assuming Gaussian distributions of the velocities and standard errors, appropriate standard error calculations should result in ca. 68% of the ratio population smaller than 1 (i.e., 68% of the velocity biases are included in their standard errors) and ca. 95% of the population smaller than 2 (i.e., 95% of the velocity biases are included in twice their standard errors). In our dataset, only 54% of the ratio are smaller than 1 and 75% are smaller than 2 (Fig. 8). . . .

Page 12 line 14. I did not see Masson et al., 2018 in the reference list.

> Yes, thank you.

Page 13 line 28. “A significant outcome of our analysis is the fact that very long series durations (over 15 – 20 years) do not ensure a better accuracy compare to series with 8 – 10 years of measurements”. However, Figure 4 says they are still getting more precise even at 20 years (though apparently at a decreasing rate of improvement) so I’m not sure if this statement is strictly true. It may be true that if a specific requirement for uncertainty is 0.1 mm/yr then there is no need to collect longer time series, but that requirement standard depends on the application and we may not yet know all future standards that are needed from the data.

> Yes I understand. It's not clear.

Page 14 Line 25: A significant outcome of our analysis is that, beyond 8 years of data, it is the presence of offsets and the noise level that have the greatest impact on the velocity bias, and not the lengthening of the series (within the limit of 21 years tested here). This suggests that the lengthening of the series is not a sufficient condition to significantly reduce the bias in estimated velocities (below the 0.1 mm yr-1 level). This effect derives directly from our noise model definition, in which the noise amplitude follows a linear power-law dependency on the frequency (Eq. 2). As a result, the noise amplitude constantly increases with long periods, explaining the very small effect of the time series duration past ca. 10 years (cf. Fig. 4). Alternative noise models, such as Gauss-Markov, that predicts a flattening of the power spectrum at long periods would likely change our results and reinstate a strong duration dependency for very long series. This shows the importance of better characterization of the GPS noise nature at very long periods and of current efforts to model and correct for long-period signals such as pluri-annual environmental loads.

Page 14 line 5. Acknowledgements sections often now contain proper attribution to those who collected (in this case the RENAG network), archived, processed the data, and from where the processed time series were downloaded, i.e ftp server, web site, etc., and on what date. In this case the authors may have had prior access to the data (?), i.e. processed it themselves, but it would improve repeatability of this work if others could be guided to where they could access the data.

> We added a "Data Availability section" to address this point and the next:

Page 15 Line 3 7 Data availability The synthetic datasets and statistical analyses were performed using R (R Core Team, 2016). The synthetic time series dataset is available upon request to the authors. Figure 9 was done with GMT5 (Wessel et al., 2011). RENAG RINEX GPS data are available from the RESIF-RENAG (RESIF., 2017). RE-

NAG GPS data were processed using the CCRS-PPP software (cf. Nguyen et al., 2016; Masson et al., 2018, for processing details). Acknowledgments We are grateful to Pr. Gilles Ducharme (IMAG, U. Montpellier) for his critical help with the regression tree analysis. We thank Simon Williams and William Hammond for their reviews that improved the quality of this manuscript.

Separate questions: Are the synthetic time series developed here openly available?

> Yes, cf. addition to the new “Data availability” section above

Page 25, line 5. Why not show b=10, discussed in the text, on the plot?

> Yes, I understand. It’s not clear. By decreasing b, the number of false detections explodes. By doing a test on a part of the dataset we saw that it was not useful to do it on the whole. So as this test was not done on the entire dataset we preferred not to include it in the figure. However this allowed a potential improvement of the method which I hope will be discussed in a specific article for this method of detection of offsets.

Page 26 Line 5: For reference, partial tests with b = 10 showed a dramatic increase of false detections, so we decided not to apply it to the entire dataset.

Interactive comment on Solid Earth Discuss., https://doi.org/10.5194/se-2018-77, 2018.